# 水稻重要病蟲害抗性基因組研究的生物資訊蒐集及統計

魏夢麗、呂秀英*、呂椿棠
行政院農委會農業試驗所農藝組

## 摘要

為加速水稻病蟲害抗性研究，以四個生物基因組資料庫(Oryzabase, Gramene, dbEST, KOME)的生物資訊為基礎，對水稻五種重要病蟲害(稻熱病、紋枯病、白葉枯病、褐飛蝨、白背飛蝨)的抗性基因組研究，如抗性基因、數量性狀基因座(QTL)、表現序列標幟(EST)及全長 cDNA(FL-cDNA)的註冊情形進行了初步統計，並就相關進展加以綜述討論。透過多個生物資料庫的搜尋和資料整理分析顯示，目前這五種重要水稻病蟲害抗性基因及其 QTL，均已被陸續鑑定出來，病害方面以水稻抗稻熱病基因相對較多，蟲害則以水稻抗褐飛蝨基因較多；但僅有受稻熱病及褐飛蝨誘導之水稻 EST 和 FL-cDNA 已登錄有序列資料。到 2005 年 5 月 31 日為止，NCBI 的 dbEST 資料庫中受稻熱病誘導之水稻 EST 有 85,105 條，佔其所有水稻 EST 之 22%，而褐飛蝨誘導之水稻 EST 則僅有 188 條。在 KOME 資料庫中受稻熱病、褐飛蝨誘導之水稻 FL-cDNA 各有 9、11 條，數量甚微，不到其所有水稻 FL-cDNA 之 0.04%。

**關鍵詞：** 水稻、生物資訊、基因符號、數量性狀基因座、表現序列標幟、全長 cDNA。

* 通信作者， iying@wufeng.tari.gov.tw
投稿日期：2005 年 7 月 20 日
接受日期：2005 年 9 月 23 日
作物、環境與生物資訊 2:296-307 (2005)
Crop, Environment & Bioinformatics 2:296-307 (2005)
189 Chung-Cheng Rd., Wufeng, Taichung Hsien 41301, Taiwan ROC

# Bioinformatics Collection and Statistics of Rice Disease/Pest Resistance Genomics Research

Meng-Li Wei, Hsiu-Ying Lu* and Chun-Tang Lu

Agronomy Division, Taiwan Agricultural Research Institute, Wufeng, Taichung Hsien 41301, Taiwan ROC

## ABSTRACT

To promote the development of the disease/pest resistance research in rice, a highlight of information on rice genomics was presented. We searched, collected and integrated the rice genome information of five major diseases and pests (blast fungus, sheath blight, bacterial blight, brown planthopper and white-backed planthopper) in rice from four public databases (Oryzabase, Gramene, dbEST and KOME). The data of major resistant gene, quantitative trait locus (QTL), expressed sequence tag (EST) and full length cDNA (FL-cDNA) from multiple public databases were summarized in statistics that allow the researchers to easily understand the current progress in rice disease/pest resistance genomics research. The perspective in these fields was also discussed. So far, many major genes and many QTLs in rice resistant to five diseases and pests have been detected, especially for blast fungus and brown planthopper resistance in rice. However, only ESTs and FL-cDNA in rice resistant to blast fungus and brown planthopper were found. Up to May 31, 2005, the dbEST of NCBI has recorded 85,105 rice ESTs induced by blast fungus, which occupied 22% of total rice ESTs. Only 188 rice ESTs induced by brown planthopper were found in dbEST. Only a few of FL-cDNA in rice ESTs induced by blast fungus and brown planthopper, being 9 and 11 pieces,

obtained from KOME database (less than 0.04% of rice FL-cDNA).

**Key words:** Rice (*Oryza sativa* L.), Bioinformatics, Gene symbol, Quantitative trait locus, Expressed sequence tag, Full length cDNA.

## 前言

臺灣地處亞熱帶，氣候高溫多濕，適合各種作物之栽培，病菌與害蟲亦因而容易滋生繁衍。水稻(*Oryza sativa* L.)是我國最主要糧食作物，栽培過程中常受到許多病蟲害之危害，而造成了產量損失。在臺灣，水稻常見的病蟲害包括有稻熱病、白葉枯病、紋枯病、褐飛蝨、白背飛蝨等(Rice Varieties in Taiwan 1930~1987, 1987)。這些水稻病蟲害之發生都會影響稻作產量，除了以傳統的施用藥劑和改進栽培方式之防治方法外，利用寄主植物水稻的抗性基因，選育並推廣培植抗病蟲害水稻品種是綜合防治最為經濟、有效和環保的方法之一。

傳統水稻育種的成功主要依賴於一系列優良基因(如矮化基因、抗病蟲基因和細胞質雄不稔基因等)的發掘和利用，但傳統品種改良技術是以整個染色體組導入，故有同時可能導入不良性狀的缺點，且顯性抗性基因直接的外表型選擇在當代無法區分純接合型或雜合型，須經結合後代的遺傳分析才能獲得抗性的純接合型，這樣的分析結果需要花費大量的人力和土地。功能基因組發現的新基因，將大大促進水稻新品種的選育。隨著「國際水稻基因組定序計畫」(IRGSP)於 2002 年 12 月宣佈水稻基因組定序草圖完成，對水稻抗性基因的研究精確度將大幅提昇。藉由水稻基因組序列比對分析和多態性鑑定，可發現品種之間的序列差異，而這些差異與外表型一致，利用這些差異將為分子標誌輔助育種 (molecular marker breeding)提供一個前所未有的機會。幾乎任何抗性基因都能找到緊密的分子標誌，利用緊密連鎖的分子標誌輔助選拔 (molecular marker-assisted selection, MAS)，可在植物生長的任一階段進行抗性基因鑑定且不受環境影響，因而簡化育種工作及節省時間(Zhang *et al*. 1996, Blair and McCouch 1997)。另外，利用 MAS 能有效地將基因累加性轉殖到抗性品種上，這是傳統育種方法很難做到的(Zhai and Zhu 1999)。培育抗病蟲害品種的關鍵在於掌握較好的抗性基因來源，而寄主植物對病蟲的侵害具有兩種遺傳背景不同的抗性，即分別由專一小種抗性質量遺傳主效基因及非專一小種抗性數量遺傳的微效基因組成 (Panda and Khush 1995)。很多研究發現，質量抗性易遭受病原菌新的生理小種和害蟲新的生物型所克服，缺乏穩定性和持久性；相形之下，數量抗性卻穩定和持久得多 (Panda and Khush 1995)，因此以數量性狀基因座位(quantitative trait locus, QTL)分子標誌定位和 MAS 方法，有利於水稻分子育種研究工作。

基因組研究對於鑑定基因表現，廣泛應用 expressed sequence tag (EST; Adams *et al*. 1991)、serial analysis of gene expression (SAGE; Velculescu *et al*. 1995)及 massively parallel signature sequencing (MPSS; Brenner *et al*. 2000)等分析方法來探討不同組織的基因表現。利用 SAGE 及 MPSS 兩種分析方法是研究一群基因表現在轉錄分子量上的有效工具，但是由於其建構基因庫(library)的選殖過程繁複，因此較少被應用在植物基因組研究(Gowda *et al*. 2004)。EST 是指從 cDNA 基因庫中隨機挑取選殖，對其進行一次大規模定序所獲得的部份 cDNA 的 5′或 3′端序列，長度一般為 100~500 bp 左右。

EST 定序雖然無法表現全長的基因序列，但是對特定細胞或特別環境中才有基因表現的場合，進行 EST 分析可提供基因辨識、新基因發現及基因表現差異性等研究一個經濟且有效之途徑(Adams *et al*. 1995, Jongeneel 2000, Rudd 2003)。而隨著生物技術的改進，取得全長(full-length cDNA , FL-cDNA)資料，可將基因結構對 DNA 序列的影響有著更多的解釋，包括內含子(intron)、外顯子(exon)的邊界與基因密碼區之間的排列，同時對於基因密碼的轉錄(transcription)、轉譯(translation)成蛋白質的基因功能研究，也提供了更多明確的訊息(Seki *et al*. 2002)。因此 EST 及 FL- cDNA 選殖體被公認是研究基因註解(annotation)與轉錄層次的基因功能分析之最好資料來源(Kikuchi *et al.* 2003)。

水稻作為禾穀類作物功能基因組研究的模型生物，其抗病蟲分子生物學不斷發展和成熟，不但有利於水稻抗病蟲基因工程策略之擬定，也有助於其他禾穀類作物病蟲害抗性基因組的研究進展。因此，為瞭解目前全球在水稻病蟲害抗性基因組研究進度，本研究利用網路上公開資料庫，蒐集上述五種水稻重要抗病蟲害的基因、QTL、EST 及 FL-cDNA 的生物資訊，並進行整理及統計，期能作為水稻抗病蟲之分子育種的參考，並為進一步的基因功能研究提供指導及幫助。

## 材料與方法

五種水稻重要病蟲害的感染源及主要危害部位，整理如 Table 1。稻熱病由真菌(*Magnaporthe grisea*)引起，自水稻秧苗期到成熟期的整個生育階段以及稻株不同的部位都會感染發病，是稻作病害中最重要的一種，目前世界上有栽培水稻的地域，不論熱帶、溫帶或寒帶皆受其害，也是臺灣最重要的稻作病害。紋枯病亦是由真菌(*Thanatephorus cucumeris* (Frank) Donk)感染水稻葉部及穗部，在稻作第一期作及第二期作均可發生。另一稻白葉枯病由細菌(*Xanthomonas oryzae* pv. *oryzae*)引起，屬於一種維管束病害感染，插秧不久之幼苗若受到感染致病，則會造成急性萎凋而枯死；水稻遭受白葉枯病後，一般減產 20~ 30%，嚴重時甚至無收穫。褐飛蝨(*Nilaparvata lugens* (Stål))是常見的水稻蟲害，於第二期稻作乳熟期至糊熟期間危害最烈；卵產於葉鞘脊部組織內，成、若蟲均好群集於稻株基部吸取稻液危害，稻株受害輕者莖葉黃化，重者全株枯死而造成「蝨燒」現象。其他較常見的水稻蟲害尚有白背飛蝨[*Sogatella furcofera* (Horvath)]，主要侵襲稻株基葉往上黃化，植株發育受阻，嚴重危害時亦可形成「蝨燒」， 被害部位滿佈黏狀分泌物及污煤狀物。

國際上有幾個主要的水稻公開資料庫，提供了水稻重要病蟲害抗性基因組之相關資訊，其名稱、網址、參考文獻及內容列如 Table 2。Oryzabase 是一個整合型的水稻科學資料庫，內容主要涵蓋水稻的遺傳種源、基因目錄、基因圖譜、突變體及基本知識。Gramene 資料庫則收集了所有禾本科作物的農藝性狀、基因、蛋白質、突變體、圖譜等相關資訊，並提供各種圖譜之間的對映比較，可用來預測及比對基因分子標誌、EST 及 cDNA 在染色體上的相對應位置，同時也將水稻與其他禾本科作物在染色體的相對應位置互作對映。NCBI 於 1992 年成立 dbEST 資料庫，專門蒐集 EST 序列，內含非常豐富的水稻 EST 資料可供搜尋。日本國家型水稻 FL-cDNA 定序計畫始於 2000 年，並於 2003 年建構了 KOME 資料

Table 1. The important diseases and pests of rice in Taiwan.

| Disease/pest | Pathogen | The attacked aerial parts |
|---|---|---|
| Rice blast | *Magnaporthe grisea* | seedlings, leaves, nods, panicles |
| Rice sheath blight | *Thanatephorus cucumeris* (Frank) Donk | leaves, sheaths, panicles |
| Rice bacterial leaf blight | *Xanthomonas oryzae* pv. *oryzae* | seedlings, leaves, nods, panicles |
| Rice brown planthopper | *Nilaparvata lugens* (Stal) | leaves, sheaths |
| Rice white-backed planthopper | *Sogatella furcofera* (Horvath) | leaves |

Table 2. Rice specific genome database.

| Database | Provider[x] | URL | Reference | Stored information |
|---|---|---|---|---|
| Oryzabase | NIG, Japan | http://www.shigen.nig.ac.jp/rice/oryzabase/top/top.jsp | Yamazaki *et al.* (2000) | Genetic resource stock information, gene dictionary, chromosome maps, mutant images and fundamental knowledge of rice science |
| Gramene | CSH, USA Cornell University, USA | http://www.gramene.org/ | Ware *et al.* (2002) | A resource for comparative genome analysis in grasses, including their agronomic traits, gene, protein, mutants, maps, etc. |
| dbEST | NCBI, USA | http://www.ncbi.nlm.nih.gov/dbEST/ | Bogusk *et al.* (1993) | EST |
| KOME | NIAS, Japan | http://cdna01.dna.affre.go.jp/cDNA | Kikuchi *et al.* (2003) | Full length cDNA of japonica rice |

[x] NIG: National Institute of Genetics; CSH: Cold Spring Harbor Laboratory; NCBI: National Center for Biotechnology Information; NIAS: National Institute of Agrobiological Sciences.

庫，該資料庫內收集了稉稻(japonica rice)之 FL-cDNA 的選殖體、序列、圖譜、核苷酸序列，並列出以 BLASTN 比對 GenBank 資料庫與以 BLASTX 比對 SWISS-PROT 資料庫的有關基因功能註解之結果。

本研究首先利用 Oryzabase 查詢抗性基因，並由 Gramene 獲取 QTL 相關資訊，以瞭解水稻重要病蟲害抗性基因組研究的總體情況；再利用 NCBI-dbEST 資料庫，搜尋水稻重要病蟲害抗性基因組的 EST 資訊，並進而利用 KOME 資料庫以搜尋水稻重要病蟲害抗性基因組的 FL-cDNA 研究。

## 結果

### 一、水稻重要病蟲害抗性基因研究

已知水稻抗病蟲基因是以多基因數廣泛分布在基因組中，針對五種水稻重要病蟲害，本研究在 Oryzabase 資料庫中，點選網頁上方的功能選單「Genes」後，利用性狀基因查詢(Trait Genes Search)的快速搜尋或進階搜尋，輸入水稻各病蟲害之抗性基因的關鍵字(keyword)或基因名稱(gene name)，並經人為逐筆瀏覽確認內容，則發現已鑑定出水稻抗病蟲害基因的數目及其

符號，整理如 Table 3。到 2005 年 5 月 31 日為止，目前已鑑定出的水稻抗病害基因數目以抗稻熱病基因最多(70 個)，其次為抗白葉枯病基因(32 個)，而抗紋枯病基因則最少(7 個)；蟲害方面則以抗褐飛蝨基因最多(29 個)，而抗白背飛蝨基因則也已鑑別出 18 個。必須注意的是，在利用 Oryzabase 資料庫搜尋抗稻熱病及抗白葉枯病基因時，得輸入兩種關鍵詞，才能查詢到較完整的基因符號。而無論輸入何種關鍵詞，都必須再經人為檢視以確認內容，例如在查詢抗稻熱病基因結果中，我們並未將 *MPi Z(Rb6)*納入，因它是一個 modifer for gene *Pi Z*；而在查詢抗白葉枯病基因結果中，*Xa23* 會重複出現在兩個不同關鍵詞中，故我們只計算一次；再以表中關鍵詞查詢抗褐飛蝨基因的最初結果中，會出現 *Bsv(Bs)*和 *Gsv(Gs)*兩個不屬於抗褐飛蝨的基因，也已被我們摒除了。

Table 3. Information summary of disease and pest resistance genes in rice provided by Oryzabase database (update time : May 31, 2005)

| Search key word (Gene name) | No. of Gene symbol hits | Gene symbol[x] |
|---|---|---|
| *Pyricularia oryza*e resistance<br>Blast resistance | 70 | *IPi*(t)*, *IPi3*(t)*, *Pi a, Pi b(Pi s), Pi f, Pi i, Pi is1(Rb4), Pi is2(Rb5), Pi k, Pi kur1*, Pi kur2*, Pi se1(Rb1), Pi se2(Rb2), Pi se3(Rb3), Pi sh, Pi ta (sl, Pi4a, Pi-ta), Pi z (Pi2, Pi-z, Piz), Pi*(t)*, *Pi1, Pi10*(t)*, *Pi11*(t)**(Pi zh), Pi12*(t), *Pi12*(t)*, *Pi13*(t), *Pi13*(t)*, *Pi14*(t), *Pi14*(t)*, *Pi15*(t), *Pi157*(t)*, *Pi157*, Pi16*(t), *Pi17*(t), *Pi18*(t)*, *Pi19*(t), *Pi20, pi21, Pi21*(t)*, *Pi22*(t)*, *Pi23*(t)*, *Pi3, Pi44*(t)*, *Pi5*(t), *Pi6*(t)*, *Pi62*(t)**(=Pita?), Pi67*(t)*, *Pi7*(t), *Pi8, Pi9*(t)*, *Pib1*, Pib2*, Piq1*, Piq2*, Piq3*, Piq4, Pit, Pb1, PBR*, Pbst*, Pi24, Pi24*(t)*(Pi-24*(t)*), Pi25, Pi25*(t)*(Pi-25*(t)*), Pi26*(t)*(Pi-26*(t)*), Pi27*(t)*(Pi-27*(t)*), Pi28*(t)*(Pi-28*(t)*), Pi33*(t), *Pi34, Pib, qBFR4-1*(t), *qBFR4-2*(t) |
| Sheath blight resistance | 7 | *qSB-11, qSB-2, qSB-3, qSB-7, qSB-9-1, SB-9-2, Rsb 1* |
| *Xanthomonas oryzae* resistance<br>Bacterial blight resistance | 32 | *Xa a*, Xa h*, xa k*, xa nm*(t)*, *xa*(t), *xa*(t)*, *Xa1(Xe), Xa10, Xa11(Xa pt), Xa12(Xa kg), xa13, Xa14, xa15*(t)**(xanm), xa16, Xa17, Xa18, xa19, Xa2, xa20, Xa21*, Xa23*(t)*, *xa24*(t)*, *Xa3 (Xa4b, Xa6, Xa9, Xa w), Xa4(Xa4a), xa5, Xa7, xa8, Xa21D, Xa23, Xa25*(t), *Xa26*(t), *Xa27*(t) |
| Brown planthopper resistance | 29 | *Bph*(t)*, *Bph1, Bph10*(t)*, *Bph14 (Qbp1), Bph15 (Qbp2), bph2, Bph3, bph4, bph5, Bph6, bph7, bph8, Bph9, IBph1, qBPH-11 (Qbph11), qBPH-12, qBPH-3 (Qbp1), qBPH-4 (Qbp2), qBPHR-1 (QBphr1), qBPHR-11-1 (QBphr11a), qBPHR-11-2 (QBphr11b), qBPHR-3 (QBphr3), qBPHR-5-1 (QBphr5a), qBPHR-5-2 (QBphr5b), qBPHR-8 (QBphr8), qEM-2, qEM-6, qGWL-5, qGWL-6* |
| Whitebacked planthopper resistance | 18 | *Ovc*(t), *qEM-3, qEM-6, qEM-7-1, qEM-7-2, qEM-8, qGWL-10, qGWL-6, qOVA-1-3*(t), *qOVA-5-1*(t), *qOVA-5-2*(t), *Wbph6, Wph1(Wbph), Wph2, Wph3, Wph4, Wph5, Wph6*(t) |

[x] (t): Currently used as a tentative locus designation; *: Not registered in Rice Genetic Cooperative; Gene symbol with capital and small letter represent dominant and recessive gene, respectively.

雖然 Oryzabase 資料庫也提供一些有關 QTL 之資訊，但從 Gramene 資料庫中可獲取更完整的水稻病蟲害抗性 QTL 資訊。進入 Gramene 主網頁後，不要直接利用快速搜尋，但點選網頁上方的功能選單「QTL」，進入 Gramene QTL 資料庫後，再點選網頁上方的「Simple Search」，於出現的「Search by Trait Category」清單中點選「Biotic stress」，則可以獲得一個已經整理好的目錄表格，該表格涵蓋了目前已知的 13 種水稻生物逆境性狀的名稱、符號、對應之存取編號及 QTL 數。透過直接點選 QTL 數，可再進一步連結至各 QTL 之更詳細資訊，如各 QTL 所對應到的連鎖群(linkage group)、遺傳圖譜位置等。倘若經由 Gramene 的快速搜尋功能，所得到的資訊缺乏整理，不易一目了然。到 2005 年 5 月 31 日為止，目前已定位出的水稻抗病蟲害 QTL 中以抗褐飛蝨及稻熱病之水稻 QTL 數最多，各有 93、83 個，其次為抗白葉枯病及白背飛蝨之水稻 QTL 各 36、20 個，而抗紋枯病之水稻 QTL 則僅有 19 個(Table 4)。

## 二、水稻重要病蟲害抗性基因組的 EST 研究

EST 雖然僅是 cDNA 的部份序列，但卻是最快速獲得序列資訊及提供功能性基因組研究的來源，可藉由基因註解進而瞭解抗性基因的機制。目前水稻 EST 資料主要集中在 NCBI 的 dbEST 資料庫，到 2005 年 5 月 31 日為止，在 dbEST 中已登錄有水稻 EST 序列 386,544 筆。為進而查詢出受病蟲害誘導表現之水稻 EST 的相關序列資料，可在 dbEST 查詢網頁中的搜尋行內，直接鍵入各水稻病蟲害之關鍵字(如 Table 5 所示)；但由於 dbEST 還包含了其他物種的 EST，單純以關鍵字進行查詢，有可能查詢結果也會將細菌、真菌和蟲的 EST 包含在內，因此本研究再將初步查詢結果利用人為瀏覽方式逐一檢視各 EST 所屬 cDNA 基因庫的內容，以過濾出病蟲害本身的 EST 資料，確保所搜尋到的資料都是屬於受病蟲害誘導之水稻 EST，結果整理如 Table 5。目前以受稻熱病誘導之水稻 EST 最多(85,105 筆)，佔所有水稻 EST 序列的 22.02%，其次是受褐飛蝨誘導之水稻 EST，但僅佔所有水稻 EST 序列的 0.05%；至於白葉枯病、紋枯病及白背飛蝨誘導之水稻 EST 的資料則迄今仍闕如(Table 5)。

針對 dbEST資料庫中受稻熱病誘導所產生的 85,105 條水稻 EST 序列，依其登錄時間及存取編號分項，共分屬於 14 個不同 cDNA 基因庫，如 Table 6 所示，這些 EST

Table 4. Information summary of disease and pest resistance QTL (markers) in rice provided by Gramene database (update time : May 31, 2005)

| Trait name | No. of QTL |
|---|---|
| Leaf blast disease resistance | 83 |
| Sheath blight disease resistance | 19 |
| Bacterial blight disease resistance | 36 |
| Brown planthopper resistance | 93 |
| White-backed planthopper resistance | 20 |

Table 5. The number of rice expressed sequence tags (ESTs) induced by different diseases and pests from NCBI-dbEST database (update time : May 31, 2005)

| Search key word | No. of EST | % |
|---|---|---|
| rice blast | 85,105 | 22.02 |
| rice sheath blight | 0 | |
| rice bacterial blight | 0 | |
| rice brown planthopper | 188 | 0.05 |
| rice white-backed planthopper | 0 | |
| rice | 386,544 | |

分別來自於不同的水稻品種(包括野生種)及不同的組織部位，序列長度介於 60~1,175 bp 之間，提供來源包括中國大陸、美國及韓國等不同國家。而褐飛蝨誘導之水稻 EST 資料主要是由中國大陸提供，分屬於感性水稻品種 Minghui 63 及抗性水稻品種 B5，都取自苗期的組織，序列長度介於 118~827 bp 之間(Table 7)。

## 三、水稻重要病蟲害抗性基因組的 cDNA 研究

從 KOME 資料庫中直接於搜尋行內直接輸入各水稻病蟲害之關鍵字(如 Table 8 所示)，可查詢到水稻抗病蟲基因的 FL-cDNA 相關資訊。至 2005 年 5 月 31 日為止，目前受稻熱病誘導之水稻 FL- cDNA 僅有 9 條，而褐飛蝨方面則有 11 條，其餘病蟲害則迄今尚無(Table 8)。由此看來，目前所獲取的水稻病蟲害抗性基因 FL-cDNA 數量仍然很少，佔所有水稻 FL-cDNA 不到 0.04%，此仍待各方竭力發掘。

Table 6. Information summary of rice ESTs induced by *M.grisea* from NCBI-dbEST database (update time : May 31, 2005)

| Genbank accession ID | Count | Entry created | Length(bp) | Cultivar[x] | Tissue | Country |
|---|---|---|---|---|---|---|
| BF889457 ~ BF889486 | 14 | 2001/5/1 | 173 ~ 577 | Drew (*japonica*) | leaf | USA |
| BI118615 ~ BI118864 | 250 | 2001/6/26 | 152 ~ 1,175 | Yuanfeengzao (*indica*) | leaf | USA |
| BI807369 ~ BI813841 | 6,437 | 2001/10/2,11/1 | 146 ~ 860 | *Oryza sativa* | leaf | China |
| BM037978 ~ BM039065 | 1,088 | 2001/11/6 | 60 ~ 1,086 | *Oryza sativa* | leaf | China |
| BM418528 ~ BM422311 | 3,784 | 2002/1/28 | 169 ~ 1,045 | *Oryza sativa* | leaf | China |
| AF521978 ~ AF522023 | 46 | 2002/7/30 | 60 ~ 600 | Bianyiyuexiangzhan (*indica*) | leaf | China |
| BQ905903 ~ BQ909262 | 3,360 | 2002/8/19 | 58 ~ 800 | *Oryza sativa* | leaf | China |
| BU666934 ~ BU667367 | 434 | 2002/9/30 | 48 ~ 788 | *Oryza sativa* | leaf | China |
| CB617709 ~ CB635999 | 18,241 | 2003/4/8 | 159 ~ 979 | IR 36 (*japonica*) | leaf | USA |
| CB636000 ~ CB686047 | 49,403 | 2003/4/8,4/9 | 152 ~ 973 | Nipponbare (*japonica*) | leaf | USA |
| BF145163 ~ BF108365 | 111 | 2003/12/30 | 115 ~ 798 | Minghui 63 (*indica*) | 14 different tissues | China |
| CD347671 ~ CD347779 | 109 | 2004/3/5 | 62 ~ 815 | *Oryza* minuta 101144 (wild) | leaf | Korea |
| CX727819 ~ CX728959 | 1,141 | 2005/1/21 | 152 ~ 686 | Nipponbare (*japonica*), IR 36 (*japonica*) | leaf | USA |
| DN475717 ~ DN475431 | 287 | 2005/3/9 | 143 ~ 1,733 | *Oryza sativa* (*indica* cultivar-grop) | 13 different tissues | India |
| Total | 85,105 | | 48 ~ 1,733 | | | |

[x] *Oryza sativa*: variety name wasn't described by submitter .

Table 7. Information summary of rice ESTs induced by brown planthopper from NCBI-dbEST database (update time : May 31, 2005)

| Genbank accession ID | Count | Entry created | Length(bp) | Cultivar[x] | Tissue | Country |
|---|---|---|---|---|---|---|
| BQ548081~BQ548082<br>BQ529294~BQ529324 | 31 | 2002/6/19 | 135 ~ 588 | Minghui 63 | seedling | China |
| BQ548083~BQ548087<br>BQ579829 | 6 | 2002/6/19 | 237 ~ 554 | B5 | seedling | China |
| CF568763~CF568780 | 18 | 2003/9/23 | 118 ~ 612 | Minghui 63 | seedling | China |
| BU572309~BU572426 | 118 | 2002/9/26 | 164 ~ 827 | B5 | seedling | China |
| BU673916~BU673930 | 15 | 2002/10/26 | 231 ~ 702 | Minghui 63 | seedling | China |
| Total | 188 | | 118 ~ 827 | | | |

[x] Minghui 63: susceptible variety; B5: resistant variety

Table 8. The numbers of rice FL-cDNA induced by different diseases and pests from KOME database (update time : May 31, 2005)

| Search key word | No. of FL-cDNA | % |
|---|---|---|
| rice blast | 9 | 0.028 |
| rice sheath blight | 0 | |
| rice bacterial blight | 0 | |
| rice brown planthopper | 11 | 0.034 |
| rice white-backed planthopper | 0 | |
| rice | 32,127 | |

KOME 現已收集的稻熱病誘導之水稻 FL-cDNA 的長度介於 962~2,737 bp 之間，取得來源分別為水稻的癒合組織、莖及花，其中大部分 FL-cDNA 的 ORF(open reading frame)佔整個序列長度的 55.57~81.50% (Table 9)。褐飛蝨誘導之 11 條水稻 FL-cDNA 的來源，主要取自於水稻的莖和花之組織部位，分屬於第 2、4、6 條染色體上，序列長度介於 529~2,124 bp 之間，其中 ORF 佔 FL-cDNA 序列長度的 37.57~76.71% (Table 10)。

## 討論

生物資訊學的主要研究內容是生物資料庫及生物資訊分析。隨著水稻基因組定序計畫的實施，水稻生物資料庫數量持續成長，資料庫結構更複雜，大量新的分析方法被提出和改進，大量重要基因被發現，以及大量來自基因組層次上的分析比較結果被公布。這些結果正在日益改變人類已有的一些觀念，傳統的研究方式已經來不及迅速消化新資料，有必要採取有效方法將它們進行適當的管理和維護，以便進一步分析、處理及利用，因此掌握各種生物資料庫的使用，可帶來極大的便利性和加速生命科學的研究。目前在網路上已有多個水稻生物資料庫可供利用，但因各種資料庫各具研究重點和主題，所包含及聯繫的生物資源不一，分類方法也各有千秋，往往必須透過多個生物資料庫的搜尋、整理和綜合歸納，始能獲取更為全面、宏觀及完整的資訊。有鑑於此，為加速水稻病蟲害抗性研究，本研究利用網路上公開資料庫，蒐集五種水稻重要抗病蟲害的基因、QTL、EST 及 FL-cDNA 的生物資訊，經過整理、歸類及統計，所得資料當可

Table 9. Information summary of rice FL-cDNA induced by *M.grisea* from KOME database (update time : May 31, 2005)

| Accession ID[x] | Length(bp) | Longest ORF(bp) | ORF length (%)[2] | Chromosome | Tissue |
|---|---|---|---|---|---|
| AK067925 | 2,737 | 697 ~ 2,217 | 55.57 | 6 | shoot |
| AK102473 | 1,257 | 78 ~ 1,022 | 75.18 | 1 | callus |
| AK070823 | 1,044 | 186 ~ 488 | 29.02 | 6 | flower |
| AK070644 | 2,146 | 61 ~ 1,809 | 81.50 | 5 | flower |
| AK069424 | 1,982 | 159 ~ 1,664 | 75.98 | 1 | flower |
| AK120401 | 2,323 | 585 ~ 2,000 | 60.96 | 2 | shoot |
| AK066531 | 2,452 | 245 ~ 1,987 | 71.08 | 6 | shoot |
| AK119939 | 1,154 | 119 ~ 934 | 70.71 | unknown | unknown |
| AK058381 | 962 | 115 ~ 837 | 75.16 | unknown | green shoot |
| Total | 962~2,737 | | | | |

[x] All entries created in 2002.

[y] ORF length (%): length ratio of ORF to its FL-cDNA.

Table 10. Information summary of rice FL-cDNA induced by brown planthopper from KOME database (update time : May 31, 2005)

| Accession ID[x] | Length(bp) | Longest ORF(bp) | ORF length (%)[y] | Chromosome | Tissue |
|---|---|---|---|---|---|
| AK065134 | 955 | 56 ~ 694 | 66.91 | 2 | shoot |
| AK104525 | 994 | 43 ~ 681 | 64.29 | 2 | etiolated shoot |
| AK104103 | 1,002 | 51 ~ 689 | 63.77 | 2 | etiolated shoot |
| AK063797 | 529 | 43 ~ 291 | 47.07 | 6 | etiolated shoot |
| AK104947 | 833 | 43 ~ 681 | 76.71 | 2 | etiolated shoot |
| AK104684 | 916 | 51 ~ 689 | 69.76 | 2 | green shoot |
| AK100671 | 2,124 | 971 ~ 1,768 | 37.57 | 4 | flower |
| AK071580 | 1,514 | 44 ~ 682 | 42.21 | 2 | flower |
| AK060125 | 1,332 | 179 ~ 976 | 59.91 | 4 | etiolated shoot |
| AK104121 | 1,331 | 178 ~ 975 | 59.95 | 4 | etiolated shoot |
| AK066843 | 964 | 91 ~ 753 | 68.78 | 4 | shoot |
| Total | 529~2,124 | | | | |

[x] All entries created on 2002.

[y] ORF length(%): length ratio of ORF to its FL-cDNA.

提供水稻抗病蟲分子育種的參考，同時也為本研究室未來進一步從事水稻病蟲害抗性基因組功能研究做好預前的準備工作。

水稻病蟲害抗性機制複雜且目前所獲得的研究成果仍然有限，從現行五種重要水稻病蟲害抗性基因組研究的總體情形來看，它們的抗性基因和 QTL 均已被陸續鑑定出來，其中以水稻抗稻熱病方面所累積的生物資訊較多(已鑑定出 70 個基因，及已定位出 83 個 QTL)。由於水稻已經具有完整的遺傳圖譜及物理圖譜，蒐集這些抗病蟲害基因及分子標誌與其連鎖之染色體相對位

置，可對抗性基因進行定位、選殖，進而可透過 MAS 及基因轉殖來培育水稻抗性品種。目前中國大陸已成功地應用 MAS 技術來改良水稻品種 Minghui63 的稻熱病抗性(Sheng *et al*. 2000)和水稻品系 9311 的白葉枯病抗性 (Zhou *et al*. 2003)，並培育出抗稻白葉枯病光敏核不育系 3418S (Luo *et al*. 2003)。QTL 定位與 MAS 相結合，是今後作物抗病蟲分子育種研究上的主要策略。不少研究發現，水稻對病蟲害的抗性表現受環境因素影響很大，因此在不同環境下進行的外表型性狀鑑定可能得到不同結果，要完全消除環境帶來的誤差是不可能的(Gallun and Khush 1980)；目前很多已定位的 QTL 是在溫室條件下獲得，這些被檢測到的抗性 QTL 應該都是有條件的，一旦環境條件改變，這些 QTL 是否還都能被檢測得到，是否還會檢測到新的抗性 QTL，都是一些值得深入研究的問題。有鑒於數量性狀通常表現強烈的基因-環境交感作用(G×E)，QTL 分析結果不能從一個環境條件外推應用到另一環境條件，因此最近某些學者已經開始嘗試將作物模擬與 QTL 相結合(Yin *et al*. 2000, Reymond *et al*. 2003, Qi *et al*. 2004)。作物模型內的作物生理生態模擬具有解析、預測 G×E 的潛力，但模型參數沒有包括其變化的遺傳學基礎，倘能與 QTL 分析結合起來，就可能把隨環境而異的數量性狀分解成若干與環境獨立的生理因子(模型參數)，並將這些生理因子的遺傳機制闡述清楚，從而使 G×E 的預測成為可能(Qi *et al*. 2004)。可預見之未來，這種結合作物模擬與 QTL 為基礎的綜合性育種分析方法，將成為一種新的學科交叉研究方法，且在植物育種中深具應用前景。

EST 及 FL-cDNA 選殖體被公認是研究基因註解與轉錄層次的基因功能分析之最好資料來源(Kikuchi *et al.* 2003)，隨著 EST 序列及 FL-cDNA 序列在公開資料庫中註冊且數量迅速增長，其所隱藏的基因註解及辨識等有用資訊，將可對水稻抗病蟲害的防禦機制研究更加瞭解。但由本研究目前所蒐集到的資訊得知，水稻 FL-cDNA 定序計畫的進展仍侷限於極少數資金雄厚的大型合作團隊；在日本國家型水稻 FL-cDNA 定序計畫所建構的 KOME 資料庫中，至 2005 年 5 月 31 日為止已累積有 32,127 筆資料，但僅有受稻熱病和褐飛蝨誘導之水稻 FL-cDNA 且數量極少，由此看來，目前想從水稻 FL-cDNA 獲取更多病蟲害抗性基因結構註解之資訊，實在非常有限。

EST 定序雖然無法表現全長的基因序列，但仍代表一個完整基因的一部分。從 1991 年 EST 的概念被提出以來(Adams *et al*. 1991)，EST 序列資料的累積異常迅速。由於一個基因的表達次數越多，能夠測到的相應 EST 也越多，因此巨量累積的 EST 序列是一種寶貴的生物資訊資源。藉由 EST 分析即可瞭解基因表達情況和表達豐富度，將得到的 EST 與其他基因組資料庫(如 GenBank 等)進行相似性分析、核酸或蛋白質序列的同源性比較等生物資訊學分析，就可用來快速尋找基因組中具有重要生物功能的區域或用來釣出未知功能之基因全長，即使功能未知，至少也能從中擷取序列歧異性、同源基因家族相對演化速率及不同組織不同時期的基因表達程度等相關資訊(Gassterland and Oprea 2001, Wolfsberg and Landsman 2001)，這也是近幾年來分離和選殖新基因及基因功能研究的一個行之有效的手段。目前 dbEST 資料的增長速度遠遠超過 NCBI 其他各種資料庫，已佔基因資料庫序列總數的 60%以上，其中水稻 EST 數量已累積近 40 萬筆，然而對於水稻

五個重要病蟲害抗性基因，目前僅登錄有受稻熱病和褐飛蝨誘導之水稻 EST，其中又以受稻熱病誘導之水稻 EST 最多(85,105 筆，佔 22.02%之多)。這些巨量的稻熱病誘導之水稻 EST 的來源不一，用作選殖定序的品種、發生的植體部位及其發育時期也不盡相同，因此如何發掘及探勘受稻熱病誘導之水稻 EST 所隱含的有用資訊，已成為本研究室近年來的重要研究課題之一。

## 引用文獻


Adams MD, JM Kelley, JD Gocayne, M Dubnick, MH Polymeropoulos, H Xiao, CR Merril, A Wu, B Olde, RF Moreno, *et al.* (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. **Science** 252:1651-1656.

Adams M, AR Kerlavage, RD Fleischmann, RA Fuldner, CJ Bult, NH Lee, EF Kirkness, KG Weinstock, JD Gocayne, O White, *et al.* (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. **Nature** 377:173-174.

Blair MW, SR McCouch (1997) Microsatellite and sequence-tagged site markers diagnostic for the rice bacterial leaf blight resistance gene *Xa5*. Theor. **Appl. Genet.** 95:174-184.

Boguski MS, TMJ Lowe, CM Tolstoshev (1993) dbEST-database for “expressed sequence tags”. **Nat. Genet.** 4:332-333.

Brenner S, M Johnson, J Bridgham, G Golda, DH Lloyd, D Johnson, S Luo, S McCurdy, M Foy, M Ewan, *et al.* (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. **Nat. Biotechnol.** 18:630-634.

Gallun RL, GS Khush (1980) Genetic factors affecting expression and stability of resistance. p.63-85. *In*: Breeding Plants Resistant to Insects. FG Maxwell, PR Jennings (eds.) John Wiley and Sons, NY.

Gassterland T, M Oprea (2001) Whole-genome analysis: annotations and updates. **Curr. Opin. Struc. Biol.** 11:377-381.

Gowda M, C Jantasuriyarat, R Dean, GL Wang (2004) Robust-longSAGE: a substantially improved longSAGE method for gene discovery and transcriptome analysis. **Plant Physiol.** 134:890-897.

Jongeneel CV (2000) Searching the expressed sequence tag (EST) database: planning for genes. **Briefings in Bioinformatics** 1:76-92.

Kikuchi S, K Satoh, T Nagata, N Kawagashira, K Doi, N Kishimoto, J Yazaki, M Ishikawa, H Yamada, H Ooka, *et al.* (2003) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. **Science** 301:376-379.

Luo YC, SH Wang, CQ Li, DZ Wang, S Wu, SY Du (2003) Breeding of the photoperiod-sensitive genetic male-sterile line ‘3418S’ resistant to bacterial blight in rice by molecular marker-assisted selection. (in Chinese) **Acta Agron. Sinica** 29:402-407.

Panda N, GS Khush (1995) Host Plant Resistance to Insects. CAB International, Oxon, UK. 431pp.

Qi CH, TJ Liu, JJ Tang, H Li (2004) Combining crop modeling and QTL mapping into an integral plant breeding strategy. (in Chinese) **Scientia Agric. Sinica** 37:1390-1395.

Reymond M, B Muller, A Leonardi, A Charcosset, F Tardieu (2003) Combining quantitative traits loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. **Plant Physiol.** 131:664-675.

Rice Varieties in Taiwan 1930~1987 (1987) Council of Agriculture, Executive Yuan ; Department of Agriculture & Forestry Taiwan Provincial Government ; Food & Fertilizer Technology Center Asia and Pacific Council Press, Taipei (in Chinese). p.307-312.

Rudd S (2003) Expressed sequence tags: alternative or complement to whole genome sequences? **Trends Plant Sci.** 8:321-329.

Seki M, M Narusaka, A Kamiya, J Ishida, M Satou, T Sakurai, *et al.* (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. **Science** 296:141- 145.

Sheng C, XH Lin, CG Xu, *et al*. (2000) Improvement of bacterial blight resistance of "Minghui63", an elite restorer line of hybrid rice by molecular marker-assisted selection. **Crop Sci.** 2000:234-244.

Velculescu VE, L Zhang, B Vogelstein, KW Kinzler (1995) Serial analysis of gene expression. **Science** 270:484-487.

Ware D, P Jaiswal, J Ni, X Pan, K Chang, K Clark, L Teytelman, S Schmidt, W Zhao, S Cartinhour, S McCouch, L Stein (2002) Gramene: a resource for comparative grass genomics. **Nucleic Acids Res.** 30**:** 103-105.

Wolfsberg TG, D Landsman (2001) Expressed sequence tags (ESTs). p.283-301. *In*: Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. DB Andreas, BFF Ouellette (eds.) 2nd ed. John Wiley & Sons, NY.

Yamazaki Y, A Yoshimura, Y Nagato, N Kurata (2000) Oryzabase-Integrated map and mutant database. Plant & Animal Genome VIII Abstract. p.54. Town & Country Convention Center. San Diego, California.

Yin X, SC Chasalow, CJ Dourleijn, P Stam, MJ Kropff (2000) Coupling estimated effects of QTLs for physiological traits to a crop growth model: Predicting yield variation among recombinant inbred lines in barley. **Heredity** 85:539-549.

Zhai WX, LH Zhu (1999) Bacterial blight resistance genes and their utilization in molecular breeding. (in Chinese) **Progress in Biotech.** 19:9-15.

Zhang G, ER Angeles, MLP Abenes (1996) RAPD and RFLP mapping of the bacterial blight resistance gene *Xa13* in rice. **Theor. Appl. Genet.** 93:65-70.

Zhou YF, HX Qi, BL Wan, XG Lu, Q Zhang, TM Mou (2003) Studies on increasing resistance of 9311 to bacterial blight by molecular marker assisted selection. (in Chinese) **Mol. Plant Breed.** 1:343-349.

－編輯：楊純明